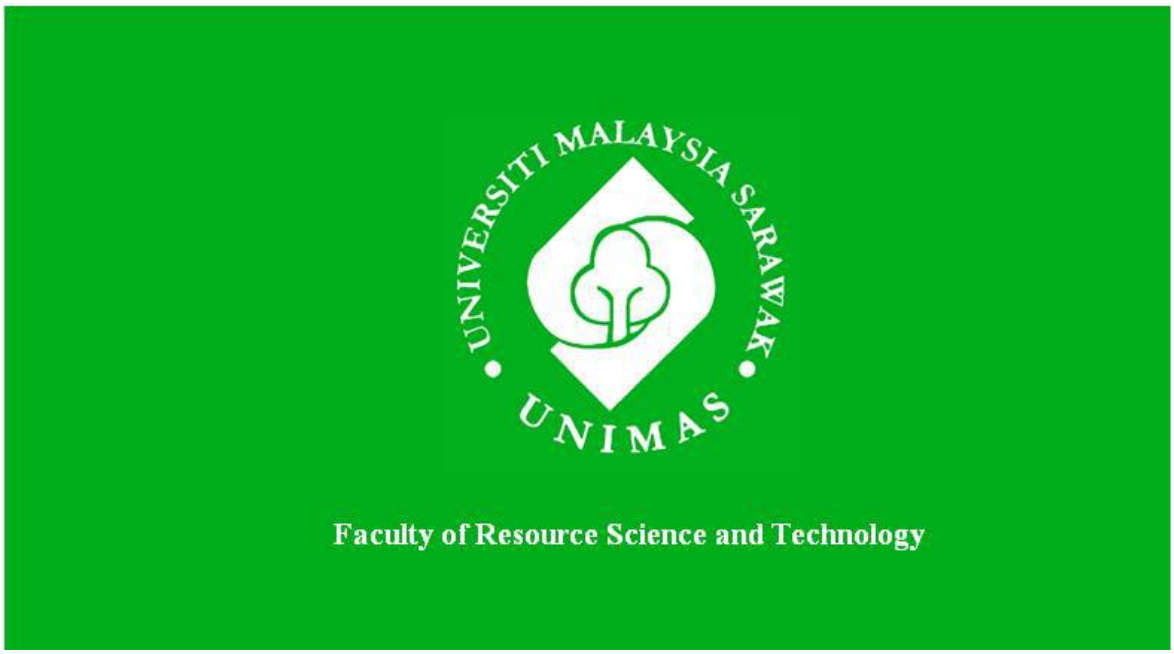

Isolation and Cloning of *ABCC2* Gene from *Rasbora Sarawakensis*

Mohd Zulfazli Bin Moktar (37148)

Bachelor of Science with Honours
(Resource Biotechnology)
2015

Isolation and Cloning of *ABCC2* Gene from *Rasbora Sarawakensis*

Mohd Zulfazli Bin Moktar (37148)

A thesis submitted in partial fulfillment of the Final Year Project 2 (STF 3015) Resource Biotechnology

Supervisor: Dr. Chung Hung Hui

Resource Biotechnology
Department of Molecular Biology

Faculty of Resource Science and Technology
Universiti Malaysia Sarawak
12/5/2015

Acknowledgement

First of all, Alhamdulillah, greatest gratitude to Allah my God whom always delivered and secured me in all circumstances and never forsake me in the time of needs. I would like to extend my gratitude to my supervisor, Dr. Chung Hung Hui for his guidance and support during the experimental work and even in the writing of thesis. Thank you Dr. for your continuous support, advices, encouragement and the patience as you bear with me throughout this experiment.

Special thanks to my fellow lab mates, Yen^2, Annie, Qawiem, Aimi, Azllan, Zain, Biha, Afuan, and Numan for always support and assist me along the experiment. Thanks to all of you, for the knowledge sharing and companionship for the past two semesters.

I would also like to thank Dr. Lee Kui Soon, Dr. Lesley, Dr. Awang Ahmad Sallehin, Dr. Hasnain and Dr. Hairul for allowed me to use some apparatus in their laboratory. My gratitude to Shek li and Syamil, Postgrad Animal Biotech Lab, for their guidance and advices regarding to my project.

Finally, my very special thanks to my precious and lovely family, my father Moktar Bin Abdul Rahim, my mother Sitti Naim Bintie Ahmad, my sisters Surinah Halin, Suriza Halin, Suhaila Halin and my brother Muhammad fazrul, for their continuous love, prayer, support and understanding throughout my study life.

Declaration

I hereby declare this thesis is based on my original work except for quotations and citations, which have been duty acknowledged. I also declare that it has not been previously or concurrently submitted to any other degrees at UNIMAS or other institutions.

Mohd Zulfazli Bin Moktar

Resource Biotechnology Programme

Department of Molecular Biology

Faculty of Resource Science and Technology

Universiti Malaysia Sarawak

Table of Contents

List of Abbreviations

AAV	Adeno-associated virus
ABC	ATP-Binding Cassette
ATP	Adenosine Triphosphate
BCRP	Breast Resistance Associated Protein
CES	Carboxylesterases
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
DNA	Deoxyribonucleic Acid
GSH	Glutathione
HDL	High Density Lipoprotein
MDR1	Multi-Drug Resistance Gene
mRNA	Messenger Ribonucleic Acid
MRP	Multidrug Resistance-associated Proteins
MXR	Mitoxantrone-Resistance Protein
NAPQI	N-acetyl-p-benzoquinone imine
NBD	Nucleotide-Binding Domain
NBF	Nucleotide binding folds
NCBI	National Center for Biotechnology Information
NTP	Nucleoside Triphosphate
PCR	Polymerase Chain Reaction
PFIC	Progressive Familial Intrahepatic Cholestasis
Pgp	Permeability glycoprotein
PXE	Pseudoxanthoma Elasticum
RNAi	RNA interference
rRNA	Ribosomal Ribonucleic Acid
siRNAs	small interfering RNAs
shRNA	short hairpin RNA
TMD	Transmembrane Domain
TOP I	Topoisomerase I
UGT	Uridine diphosphate glycosyltransferase
UV	Ultraviolet

LIST OF FIGURES

List of Tables

Isolation and Cloning of *ABCC2* Gene from *Rasbora Sarawakensis*

Mohd Zulfazli Bin Moktar (37148)
Resource Biotechnology
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

Abstract

ABC genes encode ABC transporter proteins which have a significant function in transporting molecules across the cells. *ABCC2* gene involve in the multidrug resistance-associated protein mechanism where loss of such gene has been implicated with the Dubin-Johnson syndrome (DJS) which caused by impaired of methotrexate elimination. The purpose of this study is to identify the expression of *ABCC2* gene in *Rasbora sarawakensis*, then to clone it into pGEM-T easy vector. Total RNA was extracted from whole fish homogenate via Tri reagent and phenol chloroform precipitation. The cDNA generated from reverse transcription and was amplified with PCR using degenerate primers targeting the conserved region of the gene. The PCR produced an approximately 696 bp amplicon which then was cloned into pGEM-T easy vector. Transformation was performed using in house prepared *E.coli* XL1-blue competent cells which produced an efficiency of 1.28×10^6 transformants/µg. Afterward, white colonies run on colony PCR and its revealed the presence of insert. Moreover, further confirmation was conducted through *NotI* restriction digestion which shown two discreet bands. Subsequently, the plasmid that was obtained from plasmid mini preparation was sent for sequencing and the result was corroborated by using BLAST which then show the highest similarity with *D. rerio ABCC2* transcript. Based on this study, the future expression identification and functional analysis of this gene in multixenobiotic mechanism can be carried out, thus establishing the *R. sarawakensis* as the ecotoxicology model for studying water condition in Sarawak.

Keywords: ABC transporter, *ABCC2*, Cloning, *Rasbora sarawakensis*, PCR.

Abstrak

ABC gen mengekod ABC pengangkut protein yang berfungsi mengangkut molekul melalui sel. ABCC2 gen terlibat dalam mekanisma penghalang pelbagai dadah protein, kehilangan gen ini menyebabkan sindrom Dubin-Johnson (DJS) yang disebabkan oleh penyingkiran methotrexate terjejas. Tujuan kajian ini adalah untuk mengenal pasti ekspresi ABCC2 gen dalam Rasbora sarawakensis dan mengklon ke pGEM-t easy vector. RNA telah diekstrak daripada homogenate ikan menggunakan TRI reagen dan mendapan fenol kloroform. Kemudian, cDNA yang dihasilkan melalui transkripsi terbalik melalui PCR bersama dengan degenerasi primer. PCR menghasilkan amplicon 696 bp yg kemudiannya diklonkan ke pGEM-T easy vector. Transformasi dilakukan dengan menggunakan kompeten sel E.coli XL1-blue yang menghasilkan kecekapan 1.28 × 10 ^ 6 transformants / µg. Selepas itu, koloni putih digandakan melalui kolony PCR untuk mengesan kehadiran gen. Pencernaan restriksi juga turut dilakukan yang menghasilkan dua band menunjukkan kehadiran gen yang diingini. Selepas itu, plasmid yang diperoleh dihantar untuk proses penjujukan dan keputusannya dianalisi melalui BLAST menunjukkan persamaan yang tinggi dengan Danio rerio ABCC2 gen. Berdasarkan kajian ini, analisis ekspresi and fungsi gen boleh dilanjutkan dalam mekanisma multixenobiotic, dan menjadikan Rasbora sarawakensis sebagai model ekotoksikologi untuk mengkaji kondisi air di Sarawak.

Kata kunci: ABC pengangkut, ABCC2, cloning, Rasbora sarawakensis, PCR.

1.0 INTRODUCTION

1.1 Background

ATP-binding cassette (ABC) proteins belong to the transporter protein family which transfers substance across the cells using energy through binding of ATP. Based on previous research, it has been found that these proteins highly conserved in the vertebrates' species which involve in cellular detoxification. ABC proteins are classified into eight groups based on their nucleotide binding domain (NBD), which are group A, B, C, D, E, F, G and H. The subfamily of ABCC mostly involves in the multidrug-resistance mechanism, where can be divided into two subclasses according to the present of additional N-terminal membrane spanning domains for long subclasses and absent of it for short subclasses (Ferreira et al., 2014).

ATP-binding cassette subfamily C member 2 belongs to the family of multidrug resistance protein. Multidrug resistance protein 2 (MRP2) in human encode for the *ABCC2* gene which involves in efflux the drugs out of the cells. *ABCC2* gene usually found in the liver and small amount can be found in the kidneys, intestine and placenta. This gene functions as the transporter for various types of substances out of the cells. For instance, *ABCC2* gene efflux the bilirubin from the liver cell into the bile where missing of this gene will cause an accumulation of the level of toxicity in the cell. Whereby the previous study have shown that, knockdown of this gene in the mice might be lethal in which high level of small interfering RNA (siRNA). siRNA can interfere the expression of specific gene by breaking down the mRNA after transcription which prevent the translation to occurs. Therefore, the multidrug resistance protein 2 cannot be generated which cause an accumulation of toxic in the mice cells (Loncar et al., 2010).

Model organism that was used for this experiment was *Rasbora sarawakensis* which was native to Borneo Island. *Rasbora sarawakensis* belongs to the cyprinidae family which is the family of the fish that mostly useful in the research with many advantages such as zebrafish. Therefore, this study aims to utilize *Rasbora sarawakensis* as the ecotoxicological model in Sarawak, since *ABCC2* gene important in the multixenobiotic resistance mechanisms.

1.2 Objectives

The objectives of this study:

1. To isolate the *ABCC2* gene from *Rasbora sarawakensis*.
2. To clone *ABCC2* gene into pGEM-T easy vector and analyzed the sequencing result through nucleotide BLAST.

2.0 LITERATURE REVIEW

2.1 Model Organism for This Project

2.1.1 Cyprinidae

The family Cyprinidae is the largest of the freshwater fish family, which consists of commonly known fish such as carps and minnows. This family is in the order of cypriniformes and the total species of cyprinidae are more than 2400 species in 220 genera (Wang et al., 2012). Cyprinids can be found in the water of North America, Eurasia and Africa and this family was derived from Asia based on the fossil evidence (Nelson, 2006). The sizes in length of this family of fish usually from 12 mm and up to 2.5 m. Nevertheless, this family of fish has a Weberian organ which is set of boney ossicles that join the inner ear to the swim bladder, which can amplify sound waves and letting fishes to recognize a distant long range of auditory stimuli (Wang et al., 2012).

Cyprinids fish are really important to human where some of them are serving as the food and have been farmed from a long time ago. The cyprinids fishes are commonly reared in ponds, which function as biocontrol agents to eliminate some pests such as mosquitos. Besides that, cyprinids fishes usually use as ornamental pets which be cared in the aquarium (Thai et al., 2007).

2.1.2 Rasbora genus

Rasbora is a genus of small minnow-type fish from the family Cyprinidae. This genus of fish is small, peaceful schooling fish that are easily kept and for maintenance in the aquarium. These fish usually live in the group of six or more fish. Rasboras are suitable if maintained in the planted community aquarium, because it shows beautiful coloration, and unique color outlines. Moreover, the genus Rasbora was found mostly in Southeast

Asia and Africa while its size usually small which only up to 10 cm long (Liao, 2010). Several type species of this genus are typically reared in the aquariums such as *Rasbora borapetensis, Rasbora tuberculata, Rasbora rubrodorsalis* and *Rasbora galaxy.*

2.1.3 *Rasbora sarawakensis*

Rasbora Sarawakensis is a species of small fish that can be found in the island of Borneo. From the previous research, *Rasbora sarawakensis* was found to live in numerous river systems such as Sungai Sarawak and Batang Kayan in Sarawak. *Rasbora sarawakensis* can grow until 5 cm in length and with this smaller size; it allows this fish to be easily kept in a caged-system so that it can be made available for research purposes (Surhone et al., 2009).

Figure 2.1: *Rasbora sarawakensis* (Adapted from *Rasbora sarawakensis, 2010*)

The optimum water temperature for this fish to thrive is in between 22 °C to 26 °C, with neutral pH in the range of 6 to 7.5. Since the origin of this fish from the sluggish water, therefore the frequent filtration of the water tank is not compulsory (*Rasbora sarawakensis,* 2012). *Rasbora sarawakensis* normally feeds on small living organism such

as bloodworm and water fleas. This fish is a schooling species by nature which means it's should live with their own species, so that; this fish will develop more effectively and show its natural looking. The male of *Rasbora sarawakensis* will compete to attract female fish by displaying its best color. In comparing the gender of this fish, the female of *Rasbora sarawakensis* will have rounder-bellied and larger compare to the male (Liao, 2010).

There are several conditions in order to make the fish breed and produce more fry. Firstly is condition of the water, whereby the pH of the water should be neutral but slightly acidic and the temperature of the water is in between 22 °C to 26 °C. The hatchery system for breeding this fish must be provided in order to prevent the adult form eating the eggs. The larvae can free swim after 24 hours to 48 hours post fertilizations and the food for the young *Rasbora sarawakensis* should be *paramecium* (Tang, 2010).

2.2 ATP-Binding Cassette Transporter Protein

ATP-binding cassette (ABC) proteins are the largest family of transmembrane transporter proteins encoded in the human genome. These proteins bind to ATP to use the energy to undergo the transportation of variety of molecules through cellular membranes. The mechanism of these proteins firstly discovered from its roles as the multidrug resistance (MDR) in the chemotherapeutic treatment, which acts as the barrier for the treatment of malignant tumors in human. Furthermore, these proteins were found highly conserved in the vertebrates' species, which related to the cellular detoxification. Thus, these proteins play an important role in the multixenobiotic resistance mechanism (MXR) in order to protect the aquatic organisms from xenobiotic insults (Vasiliou et al., 2009).

The ATP-dependent protein efflux the anti-cancer drugs into the extracellular medium which pioneered the discovery of multidrug resistance (MDR) phenomenon. *ABCB1* gene was initially found to play a significant role in multidrug resistance

mechanism and subsequently being denoted as permeability glycoprotein (Pgp) (Vasiliou et al., 2009). Pgp divided into two different isoforms of the gene which is multidrug-resistance 1 (MDR1), have implicated in drug resistance and multidrug-resistance 3 (MDR2), the function has not been discovered (Loncar et al., 2010). The research on MDR abilities, which shown that high level of drugs accumulation in the cell whereas the *ABCB1* gene has been knock-out compare to the cells that transfected with *ABCB1* gene. This study proved that the significant of *ABCB1* in the transporting the drugs out of the cells. The relationship between the transporter and its substrates has showing the mutual characteristics, for instance, the small size and neutral domain, moderate hydrophobicity and include natural products, chemotherapeutic drugs. ABCC and ABCG subfamilies are another gene that plays an important role in the ecotoxicological which can work with endobiotic or xenobiotic compounds (Jeong et al., 2014).

The discovering of *ABCB1* as the multidrug resistance (MDR) was leads for the advance research on cancer cell. However, the MDR phenotype not linked with *ABCB1* expression, which then leads to discovery of ABCC subfamily. This subfamily consists of 13 members and mostly is active ATP-dependent membrane transporter for organic anions of therapeutic compounds. So far, there are five members of this subfamily involve in the multidrug-resistance mechanism which includes *ABCC1, ABCC2, ABCC3, ABCC4 and ABCC5* (Chen et al., 2003).

The members of the subfamily ABCC can be classify into two different subclasses which is “long” for the subclasses that consist additional of N-terminal membrane spanning domains and “short” subclasses for the gene that does not have it. The example of subclasses long ABCC transporter includes *ABCC1, ABCC2, ABCC3* and *ABCC6*. This subfamily protein is having about 14% to 25% of homology with subfamily B transporter (Ferreira et al., 2014). Referring to this subfamily, *ABCC1* and *ABCC2* are the only gene

that has been studying on animal models, which the protein transporter that significant in organ defense.

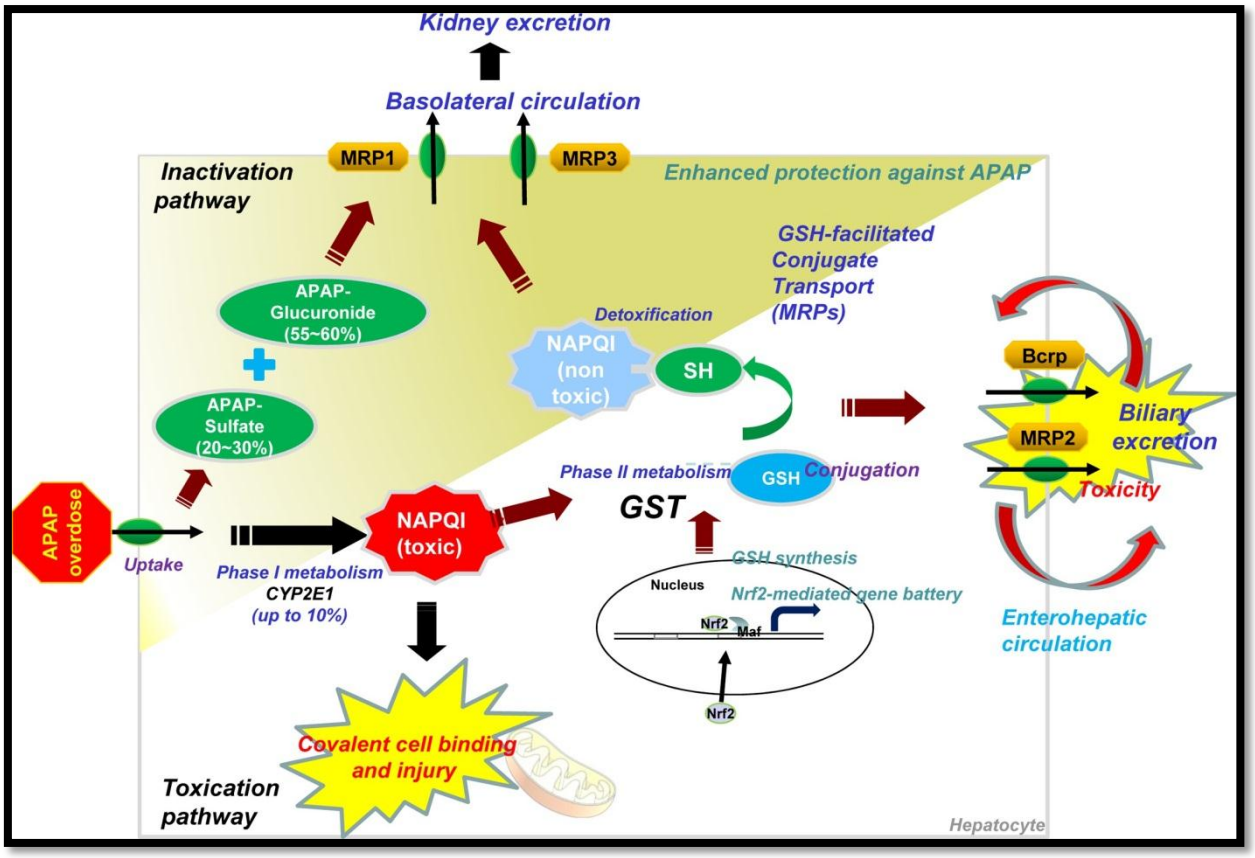

Figure 2.2: The metabolic pathways of acetaminophen then with efflux into the bile and urine in hepatocyte (Adapted from Sang & Min, 2013).

Figure above show that acetaminophen (APAP) is conjugated with glucuronic acid ensuing renal excretion. The reactive metabolite N-acetyl-p-benzoquinone imine (NAPQI) is form by metabolizing of a small fraction via CYP2E1, so that it can conjugate with sulfhydryl group of glutathione (GSH). The excretion of GSH will allows conjugation of metabolite into the bile and urine. Then, it directly initiates the enterohepatic recirculation, which will influence the progress of hepatotoxicity. Nevertheless, the unconjugated NAPQI (toxic) will covalently binds to a cellular protein under GSH diminution, which then causing the hepatocyte damaged.

2.3 ATP-binding Cassette Superfamily

The ABC proteins bind to ATP then will use the energy to transport of many molecules through the plasma membrane and intracellular membranes of the endoplasmic reticulum, peroxisome, and mitochondria. Recently, 58 members of the ABC family have been discovered, which includes 49 of human *ABC* genes and the others are from animal species where 68% of them are express in vertebrates' genomes. The subfamilies of ABC transporter are classified according to the sequence and organization of the ATP-binding domains or nucleotide binding domain (NBD), where the ABC proteins divided into eight groups (group A to H) for eukaryotes which seven (group A to G) of them are express in the human genome. The group H only has been found in the zebrafish genome with the function of this gene has not been elucidated yet (Locher et al., 2008).

ABC transporters comprise a pair of ATP-binding domains, which is nucleotide binding folds (NBF), and two sets of trans-membrane (TM) domains, usually consisting six membrane-spanning α-helices. *ABC* genes are structured whether as the full transporters which have two TM and two NBF or as the half transporters which only have one of each domain (Ferreira et al., 2014). The half transporters gather as homodimers or heterodimers to generate a functional transporter. The *ABC* genes are broadly distributed in the genome and display a high degree of amino acid sequence identity among eukaryotes. The gene superfamily can be divided into seven subfamilies and six of these subfamilies are found in both mammalian and the *S. cerevisiae* genome (Dean, 2001).

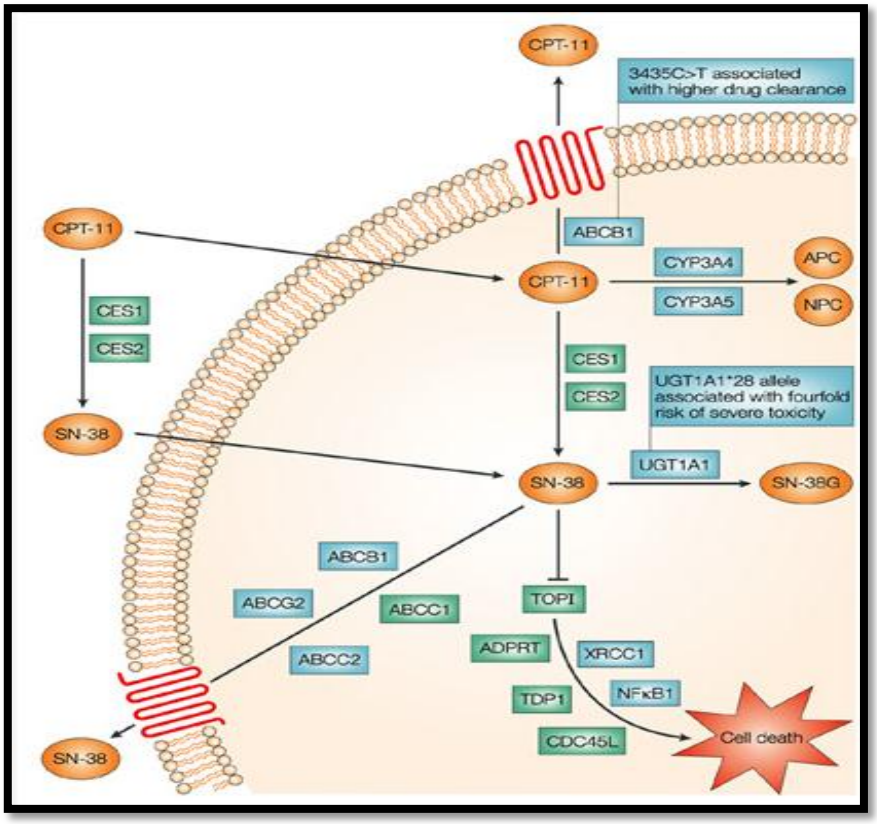

Figure 2.3: The pathways of ATP-binding cassette (Adapted from Ulrich et al., 2003).

Figure above shows the pathways of the *ABC* gene. Where the Irinotecan or Campthothecin-11 (CPT-11) can be transformed into the active metabolite SN-38 by carboxylesterases (CES) external or inside the cell. CPT-11 and SN-38 are substrates for the ATP-binding cassette (ABC) transport proteins which includes the P-glycoprotein (Pgp), ABCC and ABCG which carrying the drug go out of the cell. Otherwise, CPT-11 and SN-38 can be disabled by cytochrome P450 enzymes (CYP) or uridine diphosphate glycosyltransferase (UGT), respectively. If SN-38 continues, it will binds to topoisomerase I (TOPI), then interrupt to DNA synthesis and repair processes, which will cause the cell death (Ulrich et al., 2003).

2.4 ATP-binding Cassette subfamily C member 2

ATP-binding cassette sub-family C, member 2 or also known as the canalicular multispecific organic anion transporter 1 (cMOAT) and Multidrug resistance-associated protein 2 (MRP2) which belong to the family of multidrug resistance protein, ATP-binding cassette family and ATP gene (ATPases). This gene involve in transportation variety of substance going out of the cells. *ABCC2* gene typically found in the liver, while also can be found in the intestine, placenta and kidneys but only in a small amount. This gene has a very important role in organisms whereby it acts to remove several drugs from organ and tissue while also involve in drugs metabolism. During the drugs metabolism, this gene helps the drugs to breakdown into distinct types of component so that, the drugs will functions as required (*ABCC2,* 2014).

Figure 2.4: Location of the *ABCC2* gene in Human (Adapted from *ABCC2*, 2014)

Figure 2.4 show the location of the *ABCC2* gene in human which is in chromosome 10 on the long arm 24 (c10q24.2) as mark by the arrow. The most precise location is between 99,782,693 base pairs and 99,851,904 base pairs of chromosome 10 with size of the transcribe region is 69211 base pairs and contains 32 exons. Total amino acids that found on the *ABCC2* gene are 1545 amino acids which is a long transporter gene. This gene can found in other organism such as zebrafish in chromosome 13 and mouse in chromosome 19 (*ABCC2*, 2014).

3.0 MATERIALS AND METHODS

3.1 List of Materials

TRI reagent (Sigma, USA)

Chloroform

Isopropanol

TAE (Tris-acetate EDTA) buffer

EasyScript® Reverse Transcriptase (TransGen, China)

QIAquick ® Gel Extraction Kit (Qiagen, Germany)

0.1M CaCl2

Glycerol solution

pGEM®-T Easy Vector (Promega, USA)

LAIX (LB agar/ Ampicillin/ IPTG/ X-Gal)

QIAprep Spin Miniprep Kit (Qiagen, Germany)

3.2 Maintenance of *Rasbora sarawakensis.*

The decoration for the habitat of *Rasbora sarawakensis* are not really finicky, because this fish still show good coloration even grow up in a well-planted arrangement with a dark substrate. The tank was décor to become more natural by putting some floating plants to reduce the amount of light entering the tank. Filtration was installed to the aquarium to reduce the frequently of changing the tank water, which only being done once a week. The temperature of the water is in range of 22 °C to 26 °C, while the pH is in between 6.0 to 7.5 and with the hardness from 2 °H to 12 °H. *Rasbora sarawakensis* feed with the suitable fish food three times a day.

3.3 Total RNA Extraction of *Rasbora sarawakensis.*

The tissue was homogenized in TRI Reagent in the appropriate homogenizer. After that, the homogenate was centrifuged at 12,000 rpm for 10 minutes at 4 °C. Then, the supernatant was transferred into fresh tube and was put in the room temperature for 5 minutes. The chloroform was added into the tube that contain supernatant with amount of 0.2 ml per ml of TRI Reagent was used. Then, the sample was shaken vigorously for 15 seconds before was put in the room temperature for 5 minutes. The sample was centrifuged at 12,000 rpm for 15 minutes at 4 °C.

After that, the aqueous phase of the centrifuged mixture was transferred into the fresh tube and isopropanol was added with the amount of 0.5 ml per ml of TRI reagent was used during sample preparation and mixed. The sample then was put in the room temperature for 5 minutes before was centrifuged at 12,000 rpm for 10 minutes at 4 °C. The RNA pellet was form on the bottom of the tube, then the supernatants was removed and the RNA pellet was washed with 75% of ethanol with the amount of 1 ml per ml of TRI reagents was used during sample preparation. After that, it was centrifuged at 12,000 rpm for 5 minutes at 4 °C. Then, the ethanol was removed from the tube and the RNA pellet was dry via air drying for 10 minutes. Finally, the pellet was dissolve in the ultrapure water before store at -80 °C.

The RNA was run on the gel electrophoresis for the qualification purposes. 5 µl of the RNA and 1 µl of the 6X loading dye was load into the gel. The concentration of the gel used was 1.5% and run on the 80 V for 45 minutes. After that, the gel was undergoing post-staining with EtBr. Then, the gel was visualized under the Trans illuminator ultraviolet (UV).

For the quantification of the RNA, 999 µl of the ultrapure water was used to dissolve 1 µl of the RNA before put in the quartz cuvette. Then, the sample was spec using spectrophotometer and the reading of ratio $OD_{260/230,}$ $OD_{260/280}$ and the concentration of the RNA was taken.

3.4 Primer Design

Firstly, the clustal omega software (http://www.ebi.ac.uk/Tools/msa/clustalo/) was used for sequence alignment to find the conserve region between different species of freshwater fish where the sequences was taken from on National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) to form degenerate primer. Then, Primer3plus software (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/) was used to check whether the sequence selected suitable or not. After that, the Oligo calculate software (http://www.basic.northwestern.edu/biotools/oligocalc.html) was used to analyze all the suitable primer pairs designed for hairpin, palindromes, dimmers and melting temperature (Tm). Table 3.1 shows the degenerate code that were used to replace the non-conserved regions.

Table 3.1: Degenerate code for non-conserve region

Not conserve region	A,G	A,C	G,C	A,T	A,T,C	G,T,C	A,T,G,C	C,T	G,T	G,A,T	G,A,C
Degenerate code	R	M	S	W	H	B	N	Y	K	D	V

3.5 First-strand cDNA synthesis reaction (RT-PCR)

During cDNA synthesis, 2 µl of mRNA, 1 µl of Anchored Oligo (dT) Primer and 11 µl of ultrapure water was mix first before incubate in the thermocycler at 65 °C for 5 minutes